

# A Copula-Based Linear Model of Coregionalization for Non-Gaussian Multivariate Spatial Data

Pavel Krupskii and Marc G. Genton<sup>1</sup>

March 15, 2016

## Abstract

We propose a new copula model for replicated multivariate spatial data. Unlike classical models that assume multivariate normality of the data, the proposed copula is based on the assumption that some factors exist that affect the joint spatial dependence of all measurements of each variable as well as the joint dependence among these variables. The model is parameterized in terms of a cross-covariance function that may be chosen from the many models proposed in the literature. In addition, there are additive factors in the model that allow tail dependence and reflection asymmetry of each variable measured at different locations and of different variables to be modeled. Permutation asymmetry of different variables can also be obtained, thus providing greater flexibility. The likelihood of the model can be obtained in a simple form and therefore the likelihood estimation is quite fast. We use simulation studies to demonstrate the wide range of dependence structures that can be generated by the proposed model with different parameters. We apply the proposed model to temperature and pressure data and compare its performance with the performance of a popular model from multivariate geostatistics.

**Some key words:** Copula; Heavy tails; Spatial statistics; Tail asymmetry; Permutation asymmetry.

**Short title:** A Copula-Based Linear Model of Coregionalization

---

<sup>1</sup>CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. E-mail: pavel.krupskiy@kaust.edu.sa, marc.genton@kaust.edu.sa  
This research was supported by the King Abdullah University of Science and Technology (KAUST).

# 1 Introduction

Modeling spatial data is a challenging task as it often requires flexible, but simple and tractable models that can handle multivariate data. In many applications, models for multivariate spatial data are of interest for co-kriging; see Furrer and Genton (2011) and references therein. To set up the problem, we assume that we have a  $p$ -dimensional random field,  $Z(s) = \{Z_1(s), \dots, Z_p(s)\}^T$ , defined at locations  $s \in \mathbb{R}^d$ . In this case, several variables are observed at different locations and the main task is to model dependence both within each variable,  $Z_i(s)$ ,  $i = 1, \dots, p$  (at different locations), and between different variables. If  $Z(s)$  is a mean-zero Gaussian multivariate random field, only a covariance structure of  $Z(s)$  is needed to completely characterize the dependence structure of the field; i.e., the cross-covariance functions  $C_{i_1, i_2}(s_1, s_2) = \text{cov}\{Z_{i_1}(s_1), Z_{i_2}(s_2)\}$  need to be modeled. A popular approach is to model a matrix of correlations for each variable and cross-correlations of different variables using a multivariate covariogram and a cross-variogram (Chilés and Delfiner, 1999; Wackernagel, 2003) or using pseudo cross-variograms (Myers, 1991). The linear model of coregionalization (LMC) has been widely used to model cross-covariances; see Goulard and Voltz (1992), Wackernagel (2003), Gelfand et al. (2004), Apanasovich and Genton (2010), and the review by Genton and Kleiber (2015) and references therein.

The LMC for a stationary Gaussian process with zero mean assumes a linear structure:

$$Z_i(s) = \sum_{k=1}^r A_{ik} \tilde{Z}_k(s),$$

where  $\tilde{Z}_1, \dots, \tilde{Z}_r$  are independent stationary Gaussian processes, so that

$$C_{i_1, i_2}(s_1, s_2) = C_{i_1, i_2}(s_1 - s_2) = \sum_{k=1}^r \rho_k(s_{i_1} - s_{i_2}) A_{i_1 k} A_{i_2 k}, \quad 1 \leq i_1, i_2 \leq p,$$

where  $A$  is a  $p \times r$  full rank matrix and  $\rho_k(\cdot)$  are valid stationary correlation functions.

Different structures for the correlation functions have been proposed in the literature, such as the Matérn covariance function; see for example Gneiting (2002) and Gneiting et al. (2007) for a review of covariance functions. Parameters in the LMC can be estimated using the least squares method (Goulard and Voltz, 1992) or using maximum likelihood (Zhang, 2007).

Despite its simplicity and tractability, the LMC focuses on modeling the cross-covariance structure and may not be appropriate if the joint normality assumption is not valid for the multivariate spatial data under consideration. This can happen, for example, with data with strong joint dependence in the tails (i.e., when large/small values are simultaneously observed more often than predicted by the model), or with data with reflection asymmetry (i.e., when large values are simultaneously observed more often than small values, or vice versa).

To overcome this problem, more flexible copula-based models can be considered. Copulas have been used in a wide range of actuarial, financial and environmental studies; see Krupskii and Joe (2015a), Genest and Favre (2007), Patton (2006) and others. A copula is a multivariate cumulative distribution function with univariate uniform  $U(0, 1)$  marginals; this function can be used to link univariate marginals to construct the joint distribution function. Sklar (1959) showed that for continuous univariate distribution functions,  $F_1, \dots, F_n$ , and a continuous  $n$ -dimensional distribution function,  $F$ , there exists a unique copula function,  $C$ , such that  $F(z_1, \dots, z_n) = C\{F_1(z_1), \dots, F_n(z_n)\}$  for any  $z_1, \dots, z_n$ .

Some copula-based models have been proposed in the literature to model univariate spatial processes. One popular approach is to use vine copula models in which the joint distribution is constructed using bivariate linking copulas and from which different types of dependence structures can be obtained; see Kurowicka and Cooke (2006) and Aas et al. (2009) for details. The models using vine copulas have been applied to study climate,

geology, radiation and other spatial data; see Gräler and Pebesma (2011), Gräler (2014), Erhardt et al. (2014) and others. In these models, the dependence structure is selected based on the likelihood, making interpretability difficult. Moreover, likelihood estimation can be quite slow in high dimensions.

Other copula models for univariate spatial data include copulas parameterized in terms of pairwise dependencies, for instance a v-transformed copula of Bárdossy and Li (2008) and a chi-squared copula of Bárdossy (2006). These copulas are constructed by using a non-monotonic transformation of multivariate normal variables. As such, they cannot handle tail dependence. Moreover, the likelihood function has no simple form and obtaining parameter estimates in these models is a difficult task. Recently, Krupskii et al. (2015) proposed a factor copula model for spatial data that allows tail dependence and reflection asymmetry to be modeled. Likelihood estimates in that model can be obtained quite easily with some choice of the common factor, even if the number of locations is fairly large. However, to the best of our knowledge, flexible copula models have not yet been studied for multivariate spatial data with replicates, i.e., when there are several variables repeatedly measured at different locations.

In this paper, we extend the approach of Krupskii et al. (2015) and propose a model for stationary spatial data that combines the flexibility of a copula modeling approach, the interpretability of LMC, and the tractability of the Gaussian copula in high dimensions. The model and the corresponding copula are based on the following multivariate random process:

$$W_i(s) = Z_i(s) + \alpha_{i0}^U \mathcal{E}_0^U + \alpha_i^U \mathcal{E}_i^U - \alpha_{i0}^L \mathcal{E}_0^L - \alpha_i^L \mathcal{E}_i^L \quad (s \in \mathbb{R}^d, i = 1, \dots, p), \quad (1)$$

where  $Z_i(s)$  are cross-correlated Gaussian processes,  $\alpha_{i0}^U, \alpha_i^U, \alpha_{i0}^L, \alpha_i^L \geq 0$  for identifiability and  $\mathcal{E}_i^U, \mathcal{E}_i^L, \mathcal{E}_0^U, \mathcal{E}_0^L \sim \text{Exp}(1)$  are independent common factors with unit exponential distribution

that do not depend on the spatial location,  $s$ . This choice of the common factors allows different types of dependence structures to be generated and makes parameter estimation in this model fairly easy. The joint dependence of the Gaussian processes,  $Z_i(s)$ , can be modeled using LMC, and therefore the proposed model can be seen as an extension of LMC that allows tail dependence and asymmetric dependence to be modeled. Of course, more complex joint dependence of  $Z_i(s)$  can be considered as well; see the review by Genton and Kleiber (2015).

In our model, as well as in many other copula-based models, replicates are needed to estimate dependence parameters. With different sets of parameters in (1) models with the same covariance and cross-covariance structures but with different tail properties can be obtained. Repeated measurements of the multivariate spatial process are thus needed to estimate dependence both in the middle of the joint distribution and in its tails.

The remainder of paper is organized as follows. In Section 2, we describe the model (1) in detail and study its dependence properties. We first define the model in a general case with  $p \geq 2$ , and then, in the following sections, we provide more details about the bivariate case,  $p = 2$ . In Section 3, we conduct a small simulation study to show the flexibility of the model when data are modeled both in the middle of the distribution and in the tails. More details on the likelihood estimation and assessing the goodness of fit of the estimated models are given in Section 4. We discuss other choices of the common factors in (1) in Section 5 and also provide more details on the Pareto factors. We apply the proposed copula model to bivariate spatial data of temperature and atmospheric pressure in Oklahoma, USA, in Section 6, and we conclude with a discussion in Section 7.

## 2 Copula-Based Linear Model of Coregionalization

We use the following notation:  $\Phi(\cdot)$  is the cumulative distribution function of the univariate standard normal random variable, whereas  $\Phi_\Sigma(\cdot)$  is that of the multivariate standard normal random vector with correlation matrix  $\Sigma$ . For simplicity, in the bivariate case with  $\Sigma_{12} = \rho$ , we use the notation  $\Phi_\rho(\cdot)$ . Small symbols denote the corresponding densities.

We consider measurements of a random multivariate process by assuming that unobserved random factors exist that affect the joint dependence of all measurements of each variable as well as the joint dependence between every two variables. Specifically, we construct the corresponding copula by restricting model (1) to a finite set of locations  $s_1, \dots, s_n \in \mathbb{R}^d$ . Let  $\{(W_{1j}, \dots, W_{pj})\}_{j=1}^n$  be measurements of a  $p$ -variate spatial process that is observed at  $n$  different locations and let

$$W_{ij} = Z_{ij} + \alpha_{i0}^U \mathcal{E}_0^U + \alpha_i^U \mathcal{E}_i^U - \alpha_{i0}^L \mathcal{E}_0^L - \alpha_i^L \mathcal{E}_i^L \quad (j = 1, \dots, n, i = 1, \dots, p), \quad (2)$$

where  $\mathcal{E}_i^U, \mathcal{E}_i^L, \mathcal{E}_0^U, \mathcal{E}_0^L \sim_{\text{i.i.d.}} \text{Exp}(1)$  are exponential common factors that are independent of  $Z_{ij}$  and where  $Z = (Z_{11}, \dots, Z_{1n}, \dots, Z_{p1}, \dots, Z_{pn})^T$  has a multivariate standard normal distribution with some covariance matrix,  $\Sigma_Z$ . Exponential factors allow flexible dependence structures to be generated such that parameter estimation becomes quite fast. We discuss how other distributions of these factors affect the dependence properties of the resulting copula in Section 5. The structure of  $\Sigma_Z$  depends on the model for  $Z$ ; for example, one may use LMC. The correlation structure of  $W$  depends on that of  $Z$ . For the  $i$ -th variable,

$$\text{cor}(W_{i,j_1}, W_{i,j_2}) = \frac{\text{cor}(Z_{i,j_1}, Z_{i,j_2}) + (\alpha_{i0}^U)^2 + (\alpha_i^U)^2 + (\alpha_{i0}^L)^2 + (\alpha_i^L)^2}{1 + (\alpha_{i0}^U)^2 + (\alpha_i^U)^2 + (\alpha_{i0}^L)^2 + (\alpha_i^L)^2},$$

and for the  $i_1$ -th and  $i_2$ -th variables ( $i_1 \neq i_2$ ),

$$\text{cor}(W_{i_1,j_1}, W_{i_2,j_2}) = \frac{\text{cor}(Z_{i_1,j_1}, Z_{i_2,j_2}) + \alpha_{i_1 0}^U \alpha_{i_2 0}^U + \alpha_{i_2 0}^L \alpha_{i_1 0}^L}{[\{(\alpha_{i_1 0}^U)^2 + (\alpha_{i_1 0}^L)^2\} \{(\alpha_{i_2 0}^U)^2 + (\alpha_{i_2 0}^L)^2\}]^{1/2}}.$$

Note that  $\text{cov}(Z_{i,j_1}, Z_{i,j_2}) = 1$  implies  $\text{cov}(W_{i,j_1}, W_{i,j_2}) = 1$ ; this corresponds to perfect comonotonic dependence.

Let  $\mathbf{W} = (W_{11}, \dots, W_{1n}, \dots, W_{p1}, \dots, W_{pn})^T$  and let  $F_{n,p}^W$  and  $f_{n,p}^W$  respectively be the cumulative distribution function and probability density function of the vector  $\mathbf{W}$ . The function  $f_{n,p}^W$  can be obtained in a simple form; we provide more details for  $p = 2$  in Appendix A.1. Let  $F_{1,i}^W$  and  $f_{1,i}^W$  respectively be the cumulative distribution function and probability density function of  $W_{i1}$  ( $i = 1, \dots, p$ ). Let

$$\xi(z; \alpha_i^L, \alpha_i^U, \alpha_{i0}^L, \alpha_{i0}^U) = \frac{(\alpha_i^U)^3 \exp \{0.5/(\alpha_i^U)^2 - z/\alpha_i^U\} \Phi(z - 1/\alpha_i^U)}{\{(\alpha_{i0}^L + \alpha_i^U)(\alpha_i^L + \alpha_i^U)(\alpha_{i0}^U - \alpha_i^U)\}}.$$

One can show that

$$\begin{aligned} F_{1,i}^W(z) &= \Phi(z) + \xi(z; \alpha_i^L, \alpha_i^U, \alpha_{i0}^L, \alpha_{i0}^U) - \xi(-z; \alpha_i^U, \alpha_i^L, \alpha_{i0}^U, \alpha_{i0}^L) \\ &\quad + \xi(z; \alpha_{i0}^L, \alpha_{i0}^U, \alpha_i^L, \alpha_i^U) - \xi(-z; \alpha_{i0}^U, \alpha_{i0}^L, \alpha_i^U, \alpha_i^L). \end{aligned}$$

Because  $F_{1,i}^W(z)$  takes a simple form, the inverse function,  $(F_{1,i}^W)^{-1}(z)$ , can be easily calculated using numerical methods. Let  $\mathbf{u}_i = (u_{i1}, \dots, u_{in})^T$ ,  $0 \leq u_{ij} \leq 1$ ,  $j = 1, \dots, n$ . The copula and its density corresponding to the distribution of  $\mathbf{W}$  ( $C_{n,p}^W$  and  $c_{n,p}^W$ , respectively) can then be obtained as follows:

$$\begin{aligned} C_{n,p}^W(u_1, \dots, u_p) &= F_{n,d}^W \{ (F_{1,1}^W)^{-1}(u_1), \dots, (F_{1,p}^W)^{-1}(u_p) \}, \\ c_{n,p}^W(u_1, \dots, u_p) &= \frac{f_{n,p}^W \{ (F_{1,1}^W)^{-1}(u_1), \dots, (F_{1,p}^W)^{-1}(u_p) \}}{f_{1,1}^W \{ (F_{1,1}^W)^{-1}(u_1) \} \times \dots \times f_{1,p}^W \{ (F_{1,p}^W)^{-1}(u_p) \}}. \end{aligned} \quad (3)$$

Spatial data often have strong dependence in the tails and therefore a model that can handle strong tail dependence is necessary. One standard approach to measure tail dependence for a bivariate copula,  $C$ , is to use the lower and upper tail dependence coefficients,  $\lambda_L$  and  $\lambda_U$ , respectively:

$$\lambda_L = \lim_{q \rightarrow 0} C(q, q)/q \in [0, 1] \quad \text{and} \quad \lambda_U = \lim_{q \rightarrow 0} \bar{C}(1 - q, 1 - q)/q \in [0, 1],$$

where  $\bar{C}(u_1, u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$  is the survival copula. Copula  $C$  is said to have lower (upper) tail dependence if  $\lambda_L > 0$  ( $\lambda_U > 0$ ). For the Gaussian copula,  $\lambda_L = \lambda_U = 0$ . Models based on multivariate normality, such as the classical LMC, are therefore not suitable for modeling data with strong tail dependence.

Dependence properties of copula  $C_{n,p}^W$  depend on the choice of the parameters,  $\alpha_i^L, \alpha_i^U, \alpha_{i0}^L, \alpha_{i0}^U$  ( $i = 1, \dots, p$ ). For the  $i$ -th variable, the proposed copula simplifies to the one introduced in Krupskii et al. (2015), with the common factor  $V_0 = \alpha_{i0}^U \mathcal{E}_0^U + \alpha_i^U \mathcal{E}_i^U - \alpha_{i0}^L \mathcal{E}_0^L - \alpha_i^L \mathcal{E}_i^L$ . In particular, it follows that the bivariate copula,  $C_{2,i}^W$ , corresponding to the distribution of  $(u_{i1}, u_{i2})$ , has lower and upper tail dependence with  $\lambda_L = 2\Phi[-\{(1 - \rho_{1,2}^i)/2\}^{1/2}/\tilde{\alpha}_i^L]$  and  $\lambda_U = 2\Phi[-\{(1 - \rho_{1,2}^i)/2\}^{1/2}/\tilde{\alpha}_i^U]$ , where  $\rho_{1,2}^i = \text{cor}(Z_{i1}, Z_{i2})$ ,  $\tilde{\alpha}_i^L = \max(\alpha_i^L, \alpha_{i0}^L)$  and  $\tilde{\alpha}_i^U = \max(\alpha_i^U, \alpha_{i0}^U)$ . We now investigate dependence between two different variables. Without loss of generality, we consider copula  $C_{2,1:2}^W$ , corresponding to the distribution  $F_{2,1:2}^W$  of  $(u_{11}, u_{21})$  with  $\rho_{1,2}^{1:2} := \text{cov}(Z_{11}, Z_{21})$ .

We define  $\ell_n(x_1, x_2) := n[1 - F_{2,1:2}^W\{(F_{1,1}^W)^{-1}(1 - x_1/n), (F_{1,2}^W)^{-1}(1 - x_2/n)\}]$ . The limit  $\ell(x_1, x_2) := \lim_{n \rightarrow \infty} \ell_n(x_1, x_2)$  is called the stable upper tail dependence function of the limiting extreme value copula; see Segers (2012). We next show when copula  $C_{2,1:2}^W$  has upper tail dependence and compute the limit  $\ell(x_1, x_2)$ . For simplicity, we assume that  $\alpha_{10}^L = \alpha_{20}^L = \alpha_1^L = \alpha_2^L = 0$ . A similar result holds in the general case and for the lower tail as well; however, in the case of tail dependence, the formula for  $\ell(x_1, x_2)$  is more complicated when the coefficients  $\alpha_{10}^L, \alpha_{20}^L, \alpha_1^L$  and  $\alpha_2^L$  are nonzero.

**Proposition 1** Let  $\delta_1 = \alpha_{10}^U/\alpha_1^U$ ,  $\delta_2 = \alpha_{20}^U/\alpha_2^U$  and  $\delta_{12} = \delta_1 + \delta_2$ . Denote by  $y_i = x_i(1 - 1/\delta_i)$  and  $\delta_i^* = (\delta_i - 1)^{-1} - (\delta_{12} - 1)^{-1}$  ( $i = 1, 2$ ). If  $\min(\delta_1, \delta_2) < 1$ , copula  $C_{2,1:2}^W$  has no upper tail dependence and  $\ell(x_1, x_2) = x_1 + x_2$ . If  $\min(\delta_1, \delta_2) > 1$ , copula  $C_{2,1:2}^W$  has upper tail



dependence and, with  $\rho_{12} := \{(\alpha_{10}^U)^2 - 2\rho_{1,2}^U \alpha_{10}^U \alpha_{20}^U + (\alpha_{20}^U)^2\}^{1/2}/(\alpha_{10}^U \alpha_{20}^U)$ ,

$$\begin{aligned} \ell(x_1, x_2) = & \frac{\delta_1 y_1}{\delta_1 - 1} \Phi \left\{ \frac{\rho_{12}}{2} + \frac{\log(y_1/y_2)}{\rho_{12}} \right\} + \frac{\delta_2 y_2}{\delta_2 - 1} \Phi \left\{ \frac{\rho_{12}}{2} + \frac{\log(y_2/y_1)}{\rho_{12}} \right\} \\ & + y_2^{\delta_2} y_1^{1-\delta_2} \delta_2^* \exp \{0.5\delta_2(\delta_2 - 1)\rho_{12}^2\} \Phi \left\{ \rho_{12}(0.5 - \delta_2) + \frac{\log(y_1/y_2)}{\rho_{12}} \right\} \\ & + y_1^{\delta_1} y_2^{1-\delta_1} \delta_1^* \exp \{0.5\delta_1(\delta_1 - 1)\rho_{12}^2\} \Phi \left\{ \rho_{12}(0.5 - \delta_1) + \frac{\log(y_2/y_1)}{\rho_{12}} \right\}. \quad (4) \end{aligned}$$

The proof is given in Appendix A.2.

*Remark 1.* The limiting extreme value copula corresponding to  $C_{2,1:2}^W$  is  $\mathcal{C}_{2,1:2}^W(u_1, u_2) = \exp\{-\ell(-\log u_1, -\log u_2)\}$ . When  $\alpha_1^U = \alpha_2^U = 0$  and  $\min(\delta_1, \delta_2) > 1$ ,  $\mathcal{C}_{2,1:2}^W$  is the Hüsler-Reiss copula with parameter  $\lambda = \rho_{12}$ ; see Hüsler and Reiss (1989) for more details on the Hüsler-Reiss bivariate distribution. In the general case,  $\mathcal{C}_{2,1:2}^W$  is permutation symmetric if and only if  $\delta_1 = \delta_2$ ; that is, it is permutation symmetric if  $\alpha_1^U \alpha_{20}^U = \alpha_2^U \alpha_{10}^U$ . Permutation asymmetry of the extreme-value copula can be useful in applications for modeling data that are permutation asymmetric in the tails. We provide more details on different types of dependence structures that can be obtained in the proposed model in the next section.

### 3 Dependence Properties

In this section, we conduct some simulation studies to show different dependence structures that can be obtained in our model. For simplicity, we consider the bivariate case. In (1) with  $p = 2$  and  $d = 1$ , we assume  $\text{cov}\{Z_1(s_1), Z_2(s_2)\} = \exp(-\theta_0|s_1 - s_2|)$ ,  $\text{cov}\{Z_i(s_1), Z_i(s_2)\} = 0.5\{\exp(-\theta_i|s_1 - s_2|) + \exp(-\theta_0|s_1 - s_2|)\}$ ,  $i = 1, 2$  and  $s_1, s_2 = 1, \dots, 5$ . This covariance structure corresponds to a simple model for the bivariate process:

$$Z_i(s) = 2^{-1/2}\{Z_0(s) + Z_i^*(s)\}, \quad (i = 1, 2),$$

where  $Z_0(s)$ ,  $Z_1^*(s)$  and  $Z_2^*(s)$  are independent Gaussian processes with unit variance and exponential covariance functions. In applications, more flexible covariance structures can be selected if needed. The proposed structure is used for illustration purposes whereas similar results can also be obtained for different covariance structures.

Define  $\alpha^U = (\alpha_{10}^U, \alpha_1^U, \alpha_{20}^U, \alpha_2^U)^T$ ,  $\alpha^L = (\alpha_{10}^L, \alpha_1^L, \alpha_{20}^L, \alpha_2^L)^T$  and  $\theta = (\theta_0, \theta_1, \theta_2)^T$ . We consider three sets of dependence parameters,  $\alpha^U$ ,  $\alpha^L$  and  $\theta$ , that result in three models with very similar dependence structures within variables 1 and 2; however, cross-dependencies between variables 1 and 2 for these models are significantly different:

1.  $\alpha^U = (1.40, 0.50, 0.80, 0.00)^T$ ,  $\alpha^L = (0.80, 1.00, 0.60, 0.20)^T$ ,  $\theta = (0.75, 0.10, 0.40)^T$ ;
2.  $\alpha^U = (1.20, 1.00, 0.80, 0.20)^T$ ,  $\alpha^L = (1.00, 0.80, 0.60, 0.20)^T$ ,  $\theta = (0.75, 0.10, 0.40)^T$ ;
3.  $\alpha^U = (1.15, 1.05, 0.75, 0.50)^T$ ,  $\alpha^L = (1.10, 0.20, 0.60, 0.00)^T$ ,  $\theta = (0.75, 0.10, 0.45)^T$ .

For these sets of parameters, we define  $W_{ij}$  as in (2) with the distance between  $s_1$  and  $s_j$  equal  $|j - 1|$ ,  $i = 1, 2$ ,  $j = 1, \dots, 5$ . We assess the strength of the dependence between  $W_{11}$  and  $W_{1j}$  (dependence within the first variable), between  $W_{21}$  and  $W_{2j}$  (dependence within the second variable) and between  $W_{11}$  and  $W_{2j}$  (dependence between two variables),  $j = 1, \dots, 5$ . Specifically, we compute for these pairs of variables the Spearman's  $\rho$  and  $\lambda_L^q = \lambda_L^q(C) = C(q, q)/q$ ,  $\lambda_U^q = \lambda_U^q(C) = C_R(q, q)/q$ ,  $q = 0.01, 0.05, 0.10$ , where  $C$  is the corresponding copula for a given pair and  $C_R(u_1, u_2) = -1 + u_1 + u_2 + C(1 - u_1, 1 - u_2)$  is the reflected copula. The quantities  $\lambda_L^q$  and  $\lambda_U^q$  with small values of  $q$  are approximations of the tail dependence coefficients,  $\lambda_L$  and  $\lambda_U$ , respectively; therefore, we use them to assess the strength of the dependence in the tails. We plot these quantities to compare dependence structures generated by these three sets of parameters; see Fig. 1.

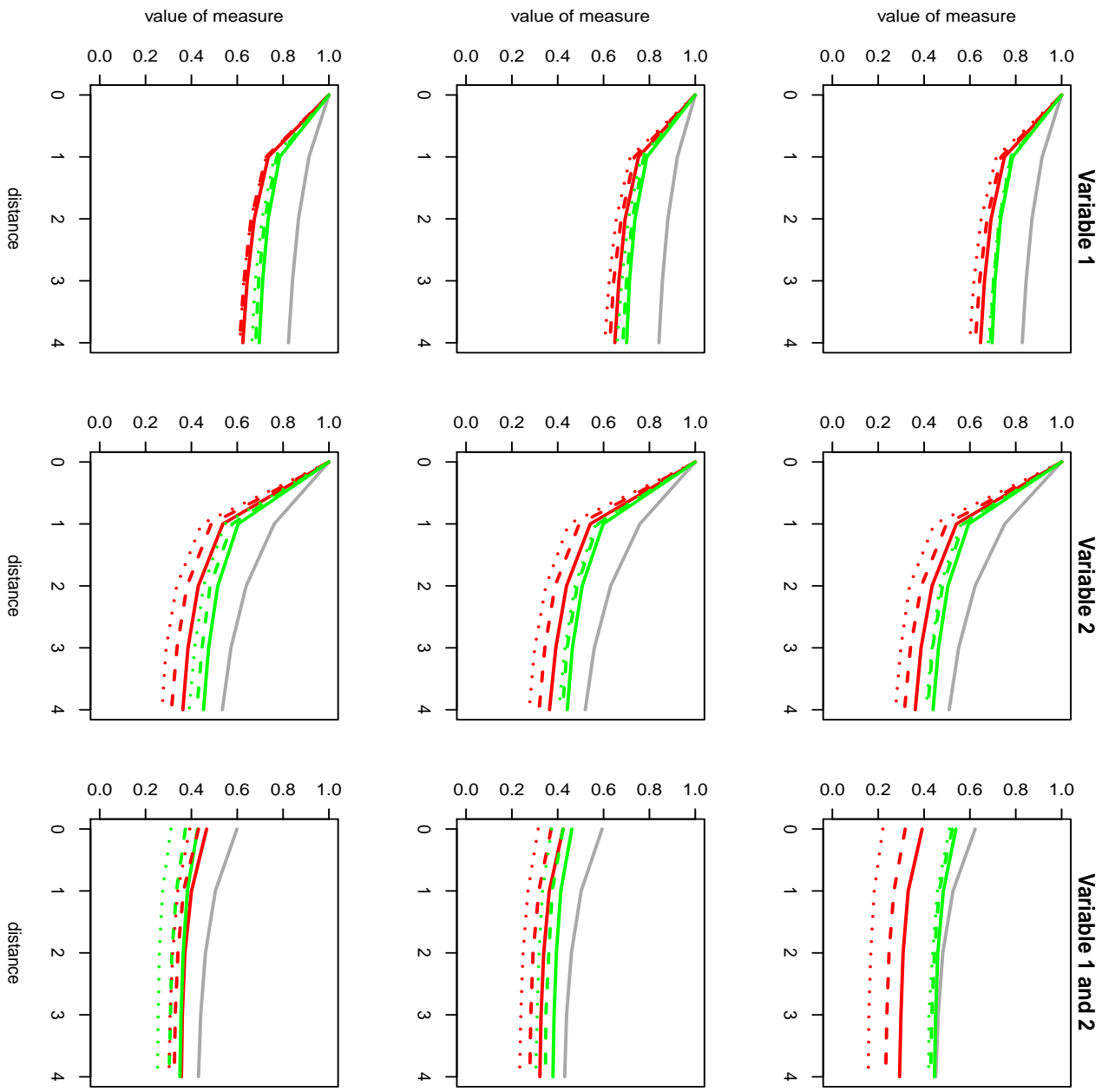


Figure 1: Spearman's  $\rho$  (grey line),  $\lambda_L^q$  (red line) and  $\lambda_U^q$  (green line), with  $q = 0.10, 0.05, 0.01$  (solid, dashed and dotted lines, respectively) for model 1 (top), model 2 (middle) and model 3 (bottom).

Because  $\theta_1 < \theta_2$  and the parameters  $\alpha_{i0}^U, \alpha_i^U, \alpha_{i0}^L, \alpha_i^L$  are larger for  $i = 1$ , the dependence within the first variable is stronger, both in the middle of the distribution and in the tails. For both variables, the dependence in the upper tail is stronger since  $\max(\alpha_{i0}^U, \alpha_i^U) > \max(\alpha_{i0}^L, \alpha_i^L)$  for  $i = 1, 2$ . In all three models, dependencies within the first and second variables are similar. This can be achieved by decreasing  $\alpha_{i0}^U$  ( $\alpha_{i0}^L$ ) and increasing  $\alpha_i^U$  ( $\alpha_i^L$ ), or vice versa. At the same time, the dependence between variables is different. For model 1,  $\alpha_{10}^L < \alpha_1^L$  and therefore copula  $C_{2,1:j}^W$ , corresponding to the distribution of  $W_{11}$  and  $W_{2j}$ , has no lower tail dependence. For models 2 and 3,  $\alpha_{i0}^L > \alpha_i^L$  and  $\alpha_{i0}^U > \alpha_i^U$  and therefore copula  $C_{2,1:j}^W$  has both lower and upper tail dependence. However, dependence in the upper tail is stronger than that in the lower tail in model 2, and it is weaker in model 3. To decrease the upper tail dependence (increase the lower tail dependence) between variables 1 and 2, one can increase the values of  $\alpha_1^U$  and  $\alpha_2^U$  (decrease the values of  $\alpha_1^L$  and  $\alpha_2^L$ ) corresponding to independent factors  $\mathcal{E}_1^U$  and  $\mathcal{E}_2^U$  ( $\mathcal{E}_1^L$  and  $\mathcal{E}_2^L$ , respectively), which are used to control the strength of the tail dependence between variables. Thus, the proposed model allows the strength of the dependence to be controlled, both in the middle of the distribution and in the tails, in each variable as well as between different variables.

## 4 Maximum Likelihood Estimation and Interpolation

### 4.1 The likelihood function

We now show how to obtain the maximum likelihood estimates for the copula parameters in model (2). We assume that we observe  $N$  independent samples,  $w_1, \dots, w_N$ , from model (2), where  $w_k = (w_{1,k}, \dots, w_{p,k})^T$ ,  $w_{i,k} = (w_{i1,k}, \dots, w_{in,k})^T$  ( $i = 1, \dots, p$ ,  $k = 1, \dots, N$ ) with essentially arbitrary marginals, not necessarily given by cdfs  $F_{1,i}^W$ . Here, vector  $w_{i,k}$  represents the  $k$ -th replicate of the  $i$ -th variable measured at  $n$  different locations. To

estimate the copula parameters, we need to transform the data to a uniform scale, e.g., non-parametrically, as follows: for each  $i = 1, \dots, p$  and  $j = 1, \dots, n$ , we can define the uniform scores,  $u_{ij,k} = \{\text{rank}(w_{ij,k}) - 0.5\}/N$  ( $k = 1, \dots, N$ ). We let  $\mathbf{z}_k = (z_{1,k}, \dots, z_{d,k})^T$ ,  $\mathbf{z}_{i,k} = (z_{i1,k}, \dots, z_{in,k})^T$ ,  $z_{ij,k} = (F_{1,i}^W)^{-1}(u_{ij,k}; \theta_{F,i})$ , where  $\theta_{F,i}$  is a vector of parameters for  $F_{1,i}^W$  ( $i = 1, \dots, p$ ,  $j = 1, \dots, n$ ,  $k = 1, \dots, N$ ). Because we use data transformed to uniform scores, they are an approximation to  $U(0, 1)$  data. Therefore, the dependence parameters can be estimated via a pseudo-likelihood function. As the number of replicates goes to infinity,  $N \rightarrow \infty$ , the likelihood estimates are consistent and asymptotically normal provided that the copula is correctly specified; see chapter 5.9 of Joe (2014) for details. Let  $\theta_F = (\theta_{F,1}, \dots, \theta_{F,p})^T$ . From (3), the pseudo log-likelihood is:

$$l(\mathbf{z}_1, \dots, \mathbf{z}_N) = \sum_{k=1}^N \log f_{n,p}^W(\mathbf{z}_{1,k}, \dots, \mathbf{z}_{p,k}; \theta_F, \theta_\Sigma) - \sum_{k=1}^N \sum_{i=1}^p \sum_{j=1}^n \log f_{1,i}^W(z_{ij,k}; \theta_{F,i}), \quad (5)$$

where  $\theta_\Sigma$  is a vector used to parameterize the correlation matrix,  $\Sigma_Z$ .

## 4.2 A conditional copula and interpolation

Let  $\hat{\theta}_F, \hat{\theta}_\Sigma$  be estimates of  $\theta_F$  and  $\theta_\Sigma$ , respectively. For a given vector of data  $(\mathbf{u}_1, \dots, \mathbf{u}_p)^T$  and the new vector (corresponding to a new,  $(n+1)$ -th, location),  $\mathbf{u}_0 = (u_{1,n+1}, \dots, u_{p,n+1})$ , on the uniform scale, where  $\mathbf{u}_i = (u_{i1}, \dots, u_{in})^T$ ,  $i = 1, \dots, p$ , we can obtain the following conditional distribution:

$$\hat{C}_{0|n,p}^W(\mathbf{u}_0 | \mathbf{u}_1, \dots, \mathbf{u}_p) := \frac{\int_0^{\mathbf{u}_0} c_{n+1,p}^W(\mathbf{u}_1^*, \dots, \mathbf{u}_p^*; \hat{\theta}_F, \hat{\theta}_\Sigma) d\mathbf{u}_0^*}{c_{n,p}^W(\mathbf{u}_1, \dots, \mathbf{u}_p; \hat{\theta}_F, \hat{\theta}_\Sigma)},$$

where  $\mathbf{u}_0^* = (u_{1,n+1}^*, \dots, u_{p,n+1}^*)^T$  and  $\mathbf{u}_i^* = (\mathbf{u}_i^T, u_{i,n+1}^*)^T$ . The conditional distribution for the  $i$ -th variable is

$$\hat{C}_{i,0|n,p}^W(u_{i,n+1} | \mathbf{u}_1, \dots, \mathbf{u}_p) = \hat{C}_{i|n,p}^W(\mathbf{u}_i | \mathbf{u}_1, \dots, \mathbf{u}_p),$$

where  $\mathbf{u}^i$  is a vector of length  $p$  such that  $(\mathbf{u}^i)_i = u_{i,n+1}$ ,  $(\mathbf{u}^i)_m = 1$  for  $m \neq i$ . Using this conditional distribution, we can calculate different quantities of interest, including the conditional expectation,  $\hat{m}_i$ , or the conditional median,  $\hat{q}_{0.5,i}$ , of the  $i$ -th variable:

$$\hat{m}_i := \int_0^1 \tilde{u}_0 \hat{c}_{i,0|n,p}^W(\tilde{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_p) d\tilde{u}_0, \quad \hat{q}_{0.5,i} := (\hat{C}_{i,0|n,p}^W)^{-1}(0.5|\mathbf{u}_1, \dots, \mathbf{u}_p),$$

where

$$\hat{c}_{i,0|n,p}^W(\tilde{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_p) = \frac{\partial \hat{C}_{i,0|n,p}^W(u_{i,n+1}|\mathbf{u}_1, \dots, \mathbf{u}_p)}{\partial u_{i,n+1}} = \frac{c_{n+1,p}^W(u_{i,n+1}, \tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_p; \hat{\theta}_F, \hat{\theta}_\Sigma)}{c_{n,p}^W(\mathbf{u}_1, \dots, \mathbf{u}_p; \hat{\theta}_F, \hat{\theta}_\Sigma)},$$

where  $\tilde{\mathbf{u}}_i = \mathbf{u}_i^*$  and  $\tilde{\mathbf{u}}_m = (\mathbf{u}_m, 1)$  for  $m \neq i$ . If  $\hat{G}_i$  is the estimated univariate marginal distribution function for the  $i$ -th variable, we can transform the uniform data to the original scale. For example, the predicted median on the original scale will be  $\hat{z}_{0.5,i} = \hat{G}_i^{-1}(\hat{q}_{0.5,i})$ . Numerical integration can be used to compute  $C_{0|n,p}^W(\mathbf{u}_0|\mathbf{u}_1, \dots, \mathbf{u}_p)$  and the inverse function  $(C_{i,0|n,p}^W)^{-1}(q|\mathbf{u}_1, \dots, \mathbf{u}_n)$  can then be used for data interpolation.

### 4.3 A simplified likelihood function for bivariate spatial data and assessing goodness of fit

In this section, we give more details about the likelihood estimation for bivariate spatial data. Consider the exponential factor model with the same covariance structure as in Section 3. Even this simple model has three covariance parameters,  $\theta_0, \theta_1, \theta_2$ , and eight dependence parameters  $\alpha_{i0}^U, \alpha_{i0}^L, \alpha_i^U, \alpha_i^L$ ,  $i = 1, 2$ . It might be difficult to estimate all these parameters, particularly if the sample size is not very large. The main reason is that different sets of these parameters may result in models with similar dependence structures. We therefore suggest that the two parameters,  $\alpha_2^U$  and  $\alpha_2^L$ , be set to zero to avoid possible overparametrization. By doing this, we avoid possible convergence problems when estimating parameters and make the estimation faster as the likelihood function simplifies in this case.

To illustrate these ideas, we consider the following bivariate model:

$$\alpha^U = (1.1, 0.9, 0.9, 0.8)^T, \alpha^L = (0.7, 1.1, 0.7, 0.4)^T, \theta = (2.2, 0.8, 1.0)^T.$$

To simulate spatial data, we use a  $5 \times 5$  grid on  $[0, 1] \times [0, 1]$  so that there are  $n = 25$  locations and  $\text{cov}(Z_{1j_1}, Z_{2j_2}) = \exp\{-\theta_0 \text{dist}(s_{1j_1}, s_{2j_2})\}$ ,  $\text{cov}(Z_{ij_1}, Z_{ij_2}) = 0.5[\exp\{-\theta_i \text{dist}(s_{ij_1}, s_{ij_2})\} + \exp\{-\theta_0 \text{dist}(s_{ij_1}, s_{ij_2})\}]$ ,  $i = 1, 2$  and  $j_1, j_2 = 1, \dots, 5$ , where  $\text{dist}(s_1, s_2)$  is the distance between location  $s_1$  and  $s_2$ . We simulate  $N = 500$  replicates from this model and then estimate parameters in this model using the constrained likelihood with  $\alpha_2^U = \alpha_2^L = 0$  (model 1). We then change the order of variables from the simulated data and again apply the restricted likelihood with  $\alpha_2^U = \alpha_2^L = 0$  (model 2). We then select the model with the best fit.

To compare the fitted dependence structure with the data, we calculated the Spearman's correlation matrices for the simulated data set and for the estimated model. We denote these matrices as  $\rho_S$  and  $\rho_S^{\text{MLE}}$  for the data and estimated model, respectively. However, the Spearman's  $\rho$  is not a good measure of dependence in the tails of a multivariate distribution. To compare the tail behavior of the two models, we therefore used the tail-weighted measures of dependence proposed by Krupskii and Joe (2015b). These measures are applied to data converted to uniform scores and defined as correlations of the transformed data. The measures provide useful summaries of the strength of the tail dependence for each pair of variables, with values close to 0 or 1 corresponding to very weak (strong, respectively) dependence in the tails. Unlike many goodness-of-fit procedures studied in the literature, the tail-weighted measures of dependence give information on how the model can be improved to fit the data in the tails better. We denoted the estimated tail-weighted measures of dependence in the lower/upper tail for the data with nonparametric estimates and for the estimated models as

$\varrho_L/\varrho_U$  and  $\alpha_L^{\text{MLE}}/\alpha_U^{\text{MLE}}$ , respectively. We obtained the model-based estimates by simulating data from the estimated models; we used the sample size  $N = 100,000$ . We calculated

$$\begin{aligned}\Delta_\rho &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n [\rho_S - \rho_S^{\text{MLE}}]_{j_1, j_2}, & |\Delta_\rho| &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n |[\rho_S - \rho_S^{\text{MLE}}]_{j_1, j_2}|, \\ \Delta_L &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n [\varrho_L - \varrho_L^{\text{MLE}}]_{j_1, j_2}, & |\Delta_L| &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n |[\varrho_L - \varrho_L^{\text{MLE}}]_{j_1, j_2}|, \\ \Delta_U &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n [\varrho_U - \varrho_U^{\text{MLE}}]_{j_1, j_2}, & |\Delta_U| &:= \frac{1}{n^2} \sum_{j_1, j_2=1}^n |[\varrho_U - \varrho_U^{\text{MLE}}]_{j_1, j_2}|.\end{aligned}$$

The quantities defined above are computed for variable 1, variable 2 (using pairs of the same variable at different locations) and variables 1 and 2 (using pairs of different variables at different locations). For comparison, we also used the further simplified model with  $\alpha_1^U = \alpha_1^L = \alpha_2^U = \alpha_2^L = 0$  (model 3). We found that model 1 fits data slightly better than does model 2, both in terms of Spearman's  $\rho$  and the tail-weighted measures of dependence,  $\alpha_L$  and  $\alpha_U$ . We therefore report the results for model 1 with the best fit. We also include model 3, which had quite a bad fit in the lower tail; see Table 1.

Table 1:  $\Delta_\rho, |\Delta_\rho|, \Delta_L, |\Delta_L|, \Delta_U, |\Delta_U|$  for models 1 and 3. Simulated data were used to calculate these values; the sample size was  $N = 100,000$ .

	$\Delta_\rho/ \Delta_\rho $	$\Delta_L/ \Delta_L $	$\Delta_U/ \Delta_U $
Model 1; loglikelihood = 28906			
Variable 1	-0.00/0.01	-0.01/0.03	-0.02/0.03
Variable 2	-0.01/0.02	-0.05/0.07	-0.02/0.03
Variables 1 and 2	0.00/0.02	-0.00/0.05	0.02/0.04
Model 3; loglikelihood = 28788			
Variable 1	-0.03/0.03	-0.04/0.04	0.02/0.03
Variable 2	-0.06/0.07	-0.15/0.15	-0.12/0.12
Variables 1 and 2	-0.07/0.07	-0.06/0.07	0.08/0.08

Model 1 fits the data quite well, including the joint dependence of each variable at different locations and the dependence between the two variables. At the same time, model



3 underestimated the dependence in the lower and upper tails for variable 2. This is because the dependence in the lower (upper) tail in variable 1, variable 2 and between variables 1 and 2 cannot be controlled when using only two dependence parameters,  $\alpha_{10}^L, \alpha_{20}^L$  ( $\alpha_{10}^U, \alpha_{20}^U$ , respectively). At least three parameters are needed to correctly specify dependencies in each tail. We obtained similar results in models with different sets of parameters and therefore we recommend using the simplified model with  $\alpha_2^U = \alpha_2^L = 0$ , which adequately assesses the strength of the dependence in bivariate spatial data.

## 5 Non-Exponential Common Factors

Possible extensions of the proposed model include models with the same structure as in (1) and (2) but with a different distribution for the variables  $\mathcal{E}_0^L, \mathcal{E}_0^U, \mathcal{E}_i^L$  and  $\mathcal{E}_i^U$ . For  $p = 1$ , Krupskii et al. (2015) showed that if this distribution has heavy tails (for example, the Pareto distribution), the resulting bivariate marginal copula for each variable will have strong tail dependence with the tail dependence coefficient equal to one. On the other hand, distributions with light tails are not suitable for modeling tail dependence.

For a non-exponential distribution of the common factors, it is difficult to obtain the joint copula density in an easy form. Multidimensional integration may be required to calculate the likelihood function. This makes parameter estimation quite complicated, especially if the number of locations and/or variables is large. Nevertheless, in this section, we provide some details for the case when  $\mathcal{E}_0^L, \mathcal{E}_0^U, \mathcal{E}_i^L, \mathcal{E}_i^U \sim_{\text{i.i.d.}} \text{Pareto}(1, k)$ , where  $k > 1$  is a shape parameter of the Pareto distribution and the scale parameter is equal to 1 (that is, the cdf  $F_{\text{Pareto}}(x) = 1 - x^{-k}$  for  $x \geq 1$ ). This case can lead to models with desirable properties in some applications. For simplicity, we again assume that  $\alpha_{10}^L = \alpha_{20}^L = \alpha_1^L = \alpha_2^L = 0$  and consider only the case of upper tail dependence.

**Proposition 2** Let  $\theta_i = (\alpha_i^U/\alpha_{i0}^U)^k$  and let  $\theta_i^* = 1/(1 + \theta_i)$ ,  $i = 1, 2$ . Then,  $\ell(x_1, x_2) = x_1 + x_2 - \min\{\theta_1^*x_1, \theta_2^*x_2\}$ . It follows that the limiting extreme value copula  $\mathcal{C}_{2,1:2}^W(u_1, u_2) = u_1u_2 \min\{u_1^{-\theta_1^*}, u_2^{-\theta_2^*}\}$ ; that is we have the Marshall-Olkin copula with parameters  $\theta_1^*$  and  $\theta_2^*$ .

The proof is given in Appendix A.3, and the properties of the Marshall-Olkin copula are discussed by Joe (2014). This result implies that copula  $\mathcal{C}_{2,1:2}^W(u_1, u_2)$  is permutation asymmetric and has tail dependence and that the tail dependence coefficient is less than one. Based on this result, we can construct a model for a univariate spatial process with a common factor and i.i.d. “Pareto noise”.

**Corollary 1** Let  $W_j = Z_j + \alpha_0^U V_0 + \alpha_1^U V_j$ ,  $j = 1, \dots, n$ , where  $Z_1, \dots, Z_n$  have multivariate normal distribution with covariance matrix  $\Sigma_Z$  and  $V_0, V_1, \dots, V_n \sim_{\text{i.i.d.}} \text{Pareto}(1, k)$ . We assume that  $V_0, V_1, \dots, V_n$  are independent of  $Z_1, \dots, Z_n$ . It follows that copula  $\mathcal{C}_2^W$ , corresponding to the distribution of  $W_1$  and  $W_2$ , has upper tail dependence and the upper tail dependence coefficient is  $\lambda_U = 1/\{1 + (\alpha_1^U/\alpha_0^U)^k\}$ .

In this model, the nugget effect is equal to  $\mathfrak{N}_0 = (\alpha_1^U)^2 \text{var}(V_1)/\{1 + (\alpha_0^U)^2 \text{var}(V_0) + (\alpha_1^U)^2 \text{var}(V_1)\} = (\alpha_1^U)^2/[(k-1)^2(1-2/k) + \{(\alpha_0^U)^2 + (\alpha_1^U)^2\}]$  for  $k > 2$ . Both  $\mathfrak{N}_0$  and  $\lambda_U$  can be controlled, depending on the choice of parameters  $\alpha_0^U, \alpha_1^U, k$ , to achieve greater flexibility in the model. The model can be readily extended to get both lower and upper tail dependence. However, as the next proposition shows, not all heavy-tailed distributions are suitable for generating flexible tail dependence similar to that in the univariate process model described above. For simplicity, we again assume that  $\alpha_{10}^L = \alpha_{20}^L = \alpha_1^L = \alpha_2^L = 0$ .

**Proposition 3** Assume that  $\mathcal{E}_0^U, \mathcal{E}_1^U, \mathcal{E}_2^U \sim_{\text{i.i.d.}} F_{\mathcal{E}}$  where  $F_{\mathcal{E}}$  belongs to the class of subexponential distributions (Chistyakov, 1964). Denote  $\bar{F}_{\mathcal{E}}(w) = 1 - F_{\mathcal{E}}(w)$ . Assume that, for

$\alpha > 1$ ,  $\bar{F}_{\mathcal{E}}(\alpha c) = o(\bar{F}_{\mathcal{E}}(c))$  as  $c \rightarrow \infty$ . If  $\alpha_{0i}^U < \alpha_i^U$  for  $i = 1$  or  $i = 2$ , then copula  $C_{2,1:2}^W(u_1, u_2)$  has no tail dependence; that is, the upper tail dependence coefficient is  $\lambda_U = 0$ . If  $\alpha_{i0}^U > \alpha_i$  for  $i = 1, 2$ , then  $\lambda_U = 1$ .

The proof is given in Appendix A.4. One example of  $F_{\mathcal{E}}$  that satisfies the assumptions of Proposition 3 is the Weibull distribution with shape parameter  $0 < \kappa < 1$ . This proposition implies that Weibull factors cannot generate flexible tail dependence between different variables and therefore these factors may be unsuitable for modeling a multivariate spatial process when the dependence between variables is generally weaker than that within each variable.

## 6 Application to temperature and atmospheric pressure data

### 6.1 Data and marginal models

In this section, we apply our model to estimate the joint dependence structure of daily mean temperature and daily mean atmospheric pressure readings in Oklahoma, USA. We consider 17 stations in the central part of the state: Acme, Apache, Chandler, Chickasha, Fort Cobb, Guthrie, Hinton, Marena, Minco, Ninnekah, OKC East, OKC North, Perkins, Shawnee, Stillwater, Washington and Watonga. These stations are located close to each other with the maximum distance between two stations being 170 kilometers. The area of interest has no big mountains and the weather conditions remain consistent across the area. The proposed copula-based LMC model 1 may therefore be suitable for modeling the joint dependence of temperature and pressure data when unobserved factors affect the temperature and pressure at all stations in this area. The data are available at [mesonet.org](http://mesonet.org).

Weather patterns can change in winter and therefore we selected observations from May

1st to September 30th, 2015, 153 days in total. To remove serial dependence, we fitted the autoregressive model with 5 lags for both temperature and pressure data. We also included a quadratic trend in the model for the temperature data, as temperatures are usually higher in July and August. We found that spatial covariates (latitude and longitude) improved the fit of the marginal model for the pressure data but not for the temperature data. One possible reason for spatial covariates to have no significant effect on the temperature data is the proximity of the weather stations and a larger variability of these data depending on the orography. We therefore can write the models for univariate marginals as follows:

$$\begin{aligned} temp_{s,t} &= \beta_0 + \beta_1 t + \beta_2 t^2 + \sum_{m=t-5}^{t-1} \alpha_{m-t+6} temp_{s,m} + \epsilon_{s,t}, \\ prss_{s,t} &= \beta_0 + \beta_1 lat_s + \beta_2 lon_s + \sum_{m=t-5}^{t-1} \alpha_{m-t+6} prss_{s,m} + \eta_{s,t}, \end{aligned}$$

where  $temp_{s,t}$  and  $prss_{s,t}$  are average temperature and pressure, respectively, measured at station  $s$  at day  $t$ ;  $lat_s$  and  $lon_s$  are spatial coordinates (latitude and longitude) of station  $s$ ,  $s = 1, \dots, 17$  and  $t = 1, \dots, 153$ . We found that the skew- $t$  distribution of Azzalini and Capitanio (2003) and the normal distribution, respectively, fit residuals  $\epsilon_{s,t}$  and  $\eta_{s,t}$  quite well. We checked the fitted residuals for uncorrelatedness using the Ljung-Box test.

## 6.2 Preliminary diagnostics of the data set

We convert the fitted residuals,  $\hat{\epsilon}_{s,t}$  and  $\hat{\eta}_{s,t}$ , from the marginal models to uniform scores. For  $s = 1, \dots, 17$ , we define

$$u_{s,t}^1 = \{\text{rank}(\hat{\epsilon}_{s,t}) - 0.5\}/153, \quad u_{s,t}^2 = \{\text{rank}(\hat{\eta}_{s,t}) - 0.5\}/153, \quad t = 1, \dots, 153.$$

If the marginal models fit the data well, the uniform scores,  $u_{s,t}^1, u_{s,t}^2$ ,  $t = 1, \dots, 153$ , should have an approximate  $U(0, 1)$  distribution for any  $s = 1, \dots, 17$ . We can therefore convert

them to the normal scores using the inverse standard normal distribution function:

$$z_{s,t}^1 = \Phi^{-1}(u_{s,t}^1), \quad z_{s,t}^2 = \Phi^{-1}(u_{s,t}^2), \quad t = 1, \dots, 153.$$

Under the assumption of joint normality, the vector  $\mathbf{z}_t = (z_{1,t}^1, \dots, z_{17,t}^1, z_{1,t}^2, \dots, z_{17,t}^2)^\top$  has a multivariate normal distribution. We can therefore draw the scatter plots for each pair of variables from the vector  $\mathbf{z}_t$  to check if these plots have the expected elliptical shape for a bivariate normal distribution. The use of the normal scores scatter plots to detect departures from normality has been advocated by Nikoloulopoulos et al. (2012). We draw the normal scores scatter plots for some pairs of the normal scores in Fig. 2.

Sharp tails in the normal scores scatter plots of the temperature data indicate that the dependence in the tails of these data is stronger than that of the normal distribution. The sharper lower tails of these scatter plots also indicate that the dependence in the temperature data might be stronger in the lower tail. The figure shows the negative dependence between the normal scores of the temperature and pressure, with a sharper lower right tail, so that the joint distribution of  $z_{s,t}^1$  and  $z_{s,t}^2$  is clearly asymmetric. Models based on the multivariate normality are therefore not suitable for modeling the joint dependence of the temperature and pressure data.

To confirm these findings, we compute the Spearman correlations and the tail-weighted dependence measures,  $\varrho_L/\varrho_U$ , from Section 4.2 for each pair of variables from the vector  $\mathbf{u}_t = (u_{1,t}^1, \dots, u_{17,t}^2, 1 - u_{1,t}^2, \dots, 1 - u_{17,t}^2)^\top$ . We use the reflected pressure data,  $1 - u_{s,t}^2$ , to get positive dependence between the temperature and pressure data. In addition, let  $\varrho_N(u_1, u_2)$  be the value  $\varrho_L(u_1, u_2) = \varrho_U(u_1, u_2)$  for data generated from the bivariate normal copula with the Spearman's  $\rho$  equal to  $\text{cor}(u_1, u_2)$ . If the bivariate copula for the pair  $(u_1, u_2)$  is a normal copula, we expect to get close values for  $\varrho_L, \varrho_U$  and  $\varrho_N$ . If the dependence in the

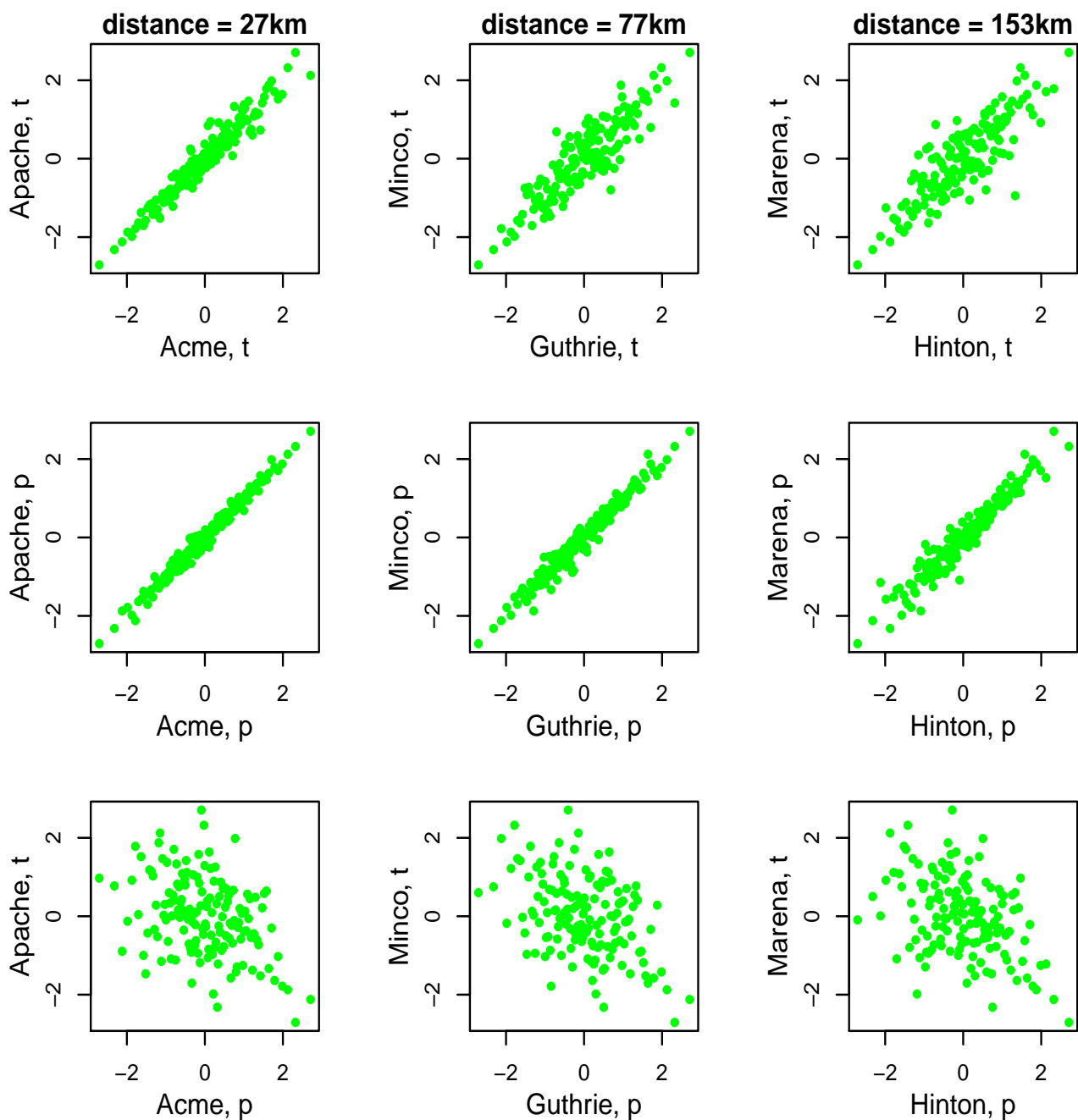


Figure 2: Scatter plots of normal scores for temperature data (top), pressure data (middle), pressure (x-axis) and temperature (y-axis) data (bottom) for Acme, Apache (left), Guthrie, Minco (middle), and Hinton, Marena (right) stations.

lower (upper) tail is stronger than that for the normal copula, then we expect the value of  $\varrho_L$  ( $\varrho_U$ ) to be larger than  $\varrho_N$ . For  $i = 1, 2$ , we compute:

$$\begin{aligned}
S_\rho^i &= \sum_{s_1 < s_2} \text{cor}(u_{s_1,t}^i, u_{s_2,t}^i)/136, & S_\rho^{12} &= \sum_{s_1 \leq s_2} \text{cor}(u_{s_1,t}^1, 1 - u_{s_2,t}^2)/153, \\
\varrho_N^i &= \sum_{s_1 < s_2} \varrho_N(u_{s_1,t}^i, u_{s_2,t}^i)/136, & \varrho_N^{12} &= \sum_{s_1 \leq s_2} \varrho_N(u_{s_1,t}^1, 1 - u_{s_2,t}^2)/153, \\
\varrho_L^i &= \sum_{s_1 < s_2} \varrho_L(u_{s_1,t}^i, u_{s_2,t}^i)/136, & \varrho_L^{12} &= \sum_{s_1 \leq s_2} \varrho_L(u_{s_1,t}^1, 1 - u_{s_2,t}^2)/153, \\
\varrho_U^i &= \sum_{s_1 < s_2} \varrho_U(u_{s_1,t}^i, u_{s_2,t}^i)/136, & \varrho_U^{12} &= \sum_{s_1 \leq s_2} \varrho_U(u_{s_1,t}^1, 1 - u_{s_2,t}^2)/153.
\end{aligned}$$

Here, the superscripts 1, 2, 12 indicate that the calculated measures are averaged for all pairs of different locations for variable 1 (temperature), variable 2 (pressure) and variables 1 and 2 (to measure cross dependencies). The results are presented in Table 2.

Table 2:  $S_\rho^i, \varrho_N^i, \varrho_L^i, \varrho_U^i$  for  $i = 1, 2$  and  $S_\rho^{12}, \varrho_N^{12}, \varrho_L^{12}, \varrho_U^{12}$

Variable	$S_\rho$	$\varrho_N$	$\varrho_L$	$\varrho_U$
Variable 1	0.85	0.74	0.86	0.79
Variable 2	0.97	0.93	0.96	0.90
Variables 1 and 2	0.38	0.20	0.55	0.17

We see that the dependence for variable 1 is stronger than it is for the normal copula in both lower and upper tails. In addition, the cross dependence between variable 1 and (reflected) variable 2 is much stronger in the upper tail. We therefore need a model that can handle tail dependence and asymmetric dependence for the temperature and pressure data.

### 6.3 Estimating the joint dependence

We apply the model (2) to the residuals obtained from the marginal models in Section 6.1 transformed to uniform scores  $u_{s,t}^1$  and  $1 - u_{s,t}^2$ ,  $s = 1, \dots, 17$  and  $t = 1, \dots, 153$ . Before fitting

the model, we need to model the cross-covariance of  $Z = (Z_{1,1}, \dots, Z_{1,17}, Z_{2,1}, \dots, Z_{2,17})^T$ .

We select the following linear model of coregionalization:

$$Z_i = \rho_i \tilde{Z}_0 + \sqrt{1 - \rho_i^2} Z_i^*, \quad i = 1, 2, \quad (6)$$

where the  $Z_0, Z_1^*, Z_2^*$  processes are independent with unit variance and have the powered exponential covariance function,  $C(d; \theta, \alpha) = \exp(-\theta d^\alpha)$  ( $\theta > 0, 0 < \alpha < 2$ ), with parameters  $(\theta_0, \alpha_0), (\theta_1, \alpha_1)$  and  $(\theta_2, \alpha_2)$ , respectively. Different models can be used to model the covariance structure of  $Z$ , including the bivariate Matérn model; we found however that these models did not improve the fit of model (6).

We set two parameters,  $\alpha_2^L, \alpha_2^U$ , to zero as discussed in Section 4.3 to avoid convergence problems with the algorithm and to increase the speed of computation. With this restriction, the maximum likelihood estimates were obtained in about five minutes on a Core i5-2410M CPU@2.3 GHz. To assess the goodness of fit of the estimated model (Model 1), we computed the values  $\Delta_\rho, \Delta_L, \Delta_U, |\Delta_\rho|, |\Delta_L|, |\Delta_U|$  defined in Section 4.3 to evaluate the strength of the dependence for the temperature and pressure data and the cross dependencies between these variables. For comparison, we also fit the LMC model (6) without exponential factors (Model 2, assuming  $\alpha_{10}^U = \alpha_1^U = \alpha_{10}^L = \alpha_1^L = \alpha_{20}^U = \alpha_2^U = \alpha_{20}^L = \alpha_2^L = 0$ ). The results are presented in Table 3.

We can see that both Model 1 and Model 2 fit the covariance structure quite well; however, Model 2 (with no exponential factors) significantly underestimates the cross dependence in the lower tail. Model 1 significantly improves the fit in the lower tail. Our proposed model is an extension of any model based on multivariate normality, including LMC (6). One can therefore apply likelihood ratio tests to check if the exponential factors are needed to model tail dependence. For example, Model 1 and Model 2 (the latter is a



Table 3:  $\Delta_\rho, |\Delta_\rho|, \Delta_L, |\Delta_L|, \Delta_U, |\Delta_U|$  for models 1 and 2. We simulated data from the estimated models 1 and 2 to calculate these values; we used  $N = 100,000$  replicates.

	$\Delta_\rho/ \Delta_\rho $	$\Delta_L/ \Delta_L $	$\Delta_U/ \Delta_U $
Model 1; loglikelihood = 8259.1			
Variable 1	0.00/0.01	-0.01/0.01	-0.02/0.02
Variable 2	0.03/0.03	-0.03/0.04	0.06/0.07
Variables 1 and 2	-0.05/0.06	0.13/0.13	0.06/0.09
Model 2; loglikelihood = 8096.5			
Variable 1	0.00/0.01	0.03/0.03	-0.03/0.03
Variable 2	-0.03/0.03	0.10/0.10	0.02/0.05
Variables 1 and 2	-0.04/0.05	0.34/0.34	-0.03/0.08

nested model) can be compared by calculating the test statistic,  $\text{LR}(\text{Model 1}, \text{Model 2}) = 2(8259.1 - 8096.5) = 325.3 > \chi^2(0.95, 6) = 12.59$ , where  $\chi^2(0.95, 6)$  is the 95% quantile of a chi-squared distribution with six degrees of freedom (Model 1 has six additional parameters,  $\alpha_{10}^U, \alpha_1^U, \alpha_{10}^L, \alpha_1^L, \alpha_{20}^U, \alpha_{20}^L$ ), suggesting that the exponential factors used to model the tail dependence significantly improve the fit. Similarly, covariance of the vector  $\mathbf{Z}$  in (2) with a different model yields two nested models (with and without exponential factors) that can be estimated and the likelihood ratio test can be applied to compare these models.

## 7 Conclusion

We proposed a new copula model for multivariate spatial data that can handle tail dependence and asymmetric dependence within each variable as well as between different variables. This model is a generalization of any model based on multivariate normality. The widely used linear model of coregionalization is an example. Parameters in the model can be estimated by likelihood and the formula for the joint copula density is quite simple so that parameters can be computed fairly easily. The model allows simple interpretation when fac-

tors affecting the joint dependence of multivariate spatial data exist. While we focused on exponential factors used to model tail dependence and asymmetry in (2), we also discussed how the choice of the distribution of these factors can affect tail properties of the resulting model.

While the proposed model can generate a wide range of dependence structures, it assumes a linear structure when some exponential factors are added to the multivariate Gaussian process. One direction for future research can therefore be to consider models with multiplicative factors or more general models based on the multivariate process:

$$W_i(s) = f_i\{Z_i(s), \mathcal{E}_0, \mathcal{E}_i\}, \quad i = 1, \dots, p, \quad (7)$$

where  $\{Z_1(s), \dots, Z_p(s)\}^T$  is a multivariate Gaussian process and  $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_p$  are factors that do not depend on the location  $s$  (not necessarily exponential), introduced to increase the flexibility of the model. The choice of functions  $f_1, \dots, f_p$  and distributions of factors  $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_p$  can define the dependence properties of the joint distribution and the resulting copula. Another direction for future research is to study anisotropic factor copula models when the joint dependence of the variables is affected by a location or some other factors. One way to introduce anisotropy can be, for example, by spatially varying functions  $f_1, \dots, f_p$  in (7).

## References

- Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44, 182–198.
- Apanasovich, T. V., Genton, M. G., 2010. Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* 97, 15–30.

- Azzalini, A., Capitanio, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society. Series B* 65, 367–389.
- Bárdossy, A., 2006. Copula-based geostatistical models for groundwater quality parameters. *Water resources research* 42.
- Bárdossy, A., Li, J., 2008. Geostatistical interpolation using copulas. *Water resources research* 44.
- Chilés, J.P., Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*. John Wiley, New York.
- Chistyakov, V.P., 1964. A theorem on sums of independent, positive random variables and its applications to branching processes. *Theory of Probability and its Applications* 9, 640–648.
- Embrechts, P., Goldie, C. M., Veraverbeke, N., 1979. Subexponentiality and infinite divisibility. *Z. Wahr. verw. Gebiete* 49, 335–347.
- Erhardt, T.M., Czado, C., Schepsmeier, U., 2014. R-vine models for spatial time series with an application to daily mean temperature. *Zentrum Mathematik Technische Universität München*.
- Feller, W., 1970. *An Introduction to Probability Theory and Its Applications*. volume Volume 2. John Wiley & Sons, USA.
- Furrer, R., Genton, M. G., 2011. Aggregation-cokriging for highly-multivariate spatial data. *Biometrika* 98, 615–631.

- Gelfand, A. E., Schmidt, A. M., Banerjee, S., Sirmans, C. F., 2004. Nonstationary multivariate process modeling through spatially varying coregionalization. *Test* 13, 263–312.
- Genest, C., Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12, 347–368.
- Genton, M. G., Kleiber, W., 2015. Cross-covariance functions for multivariate geostatistics. *Statistical Science* 30, 147–163.
- Gneiting, T., 2002. Nonseparable, stationary covariance functions for space-time data. *J. Am. Statist. Assoc.* 97, 590–600.
- Gneiting, T., Genton, M. G., Guttorp, P., 2007. Geostatistical space-time models, stationarity, separability and full symmetry. In Finkenstaedt, B., Held, L. and Isham, V.(eds), *Statistics of Spatio-Temporal Systems*, Chapman & Hall / CRC Press, Monograph in Statistics and Applied Probability, Boca Raton.
- Goulard, M., Voltz, M., 1992. Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology* 24, 269–286.
- Gräler, B., 2014. Modelling skewed spatial random fields through the spatial vine copula. *Spatial Statistics* 10, 87–102.
- Gräler, B., Pebesma, E., 2011. The pair-copula construction for spatial data: a new approach to model spatial dependency. *Procedia Environmental Sciences* 7, 206–211.
- Hüsler, J., Reiss, R. D., 1989. Maxima of normal random vectors: between independence and complete dependence. *Stat. Prob. Lett.* 7, 283–286.

- Joe, H., 2014. Dependence Modeling with Copulas. Chapman & Hall/CRC, Boca Raton, FL.
- Krupskii, P., Huser, R., Genton, M. G., 2015. Factor copula models for spatial data. <http://arxiv.org/abs/1511.03000> .
- Krupskii, P., Joe, H., 2015a. Structured factor copula models: theory, inference and computation. *Journal of Multivariate Analysis* 138, 53–73.
- Krupskii, P., Joe, H., 2015b. Tail-weighted measures of dependence. *Journal of Applied Statistics* 42, 614–629.
- Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley Series in Probability and Statistics.
- Myers, D. E., 1991. Pseudo-cross variograms, positive-definiteness, and cokriging. *Mathematical Geology* 23, 805–816.
- Nikoloulopoulos, A.K., Joe, H., Li, H., 2012. Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis* 56, 3659–3673.
- Patton, A., 2006. Modeling asymmetric exchange rate dependence. *International Economic Review* 47, 527–556.
- Segers, J., 2012. Max-stable models for multivariate extremes. *REVSTAT – Statistical Journal* 10, 61–82.
- Sklar, A., 1959. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8, 229–231.

- Stroud, A., Secrest, D., 1966. Gaussian Quadrature Formulas. Prentice-Hall, Englewood Cliffs, NJ.
- Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin. 2nd ed. edition.
- Zhang, H., 2007. Maximum-likelihood estimation for multivariate spatial linear coregionalization models. *Environmetrics* 18, 125–139.

## Appendix

### A.1 Formula for $f_{n,p}^W$ for $p = 2$

Let  $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2)^T$ , where  $\mathbf{w}_i = (w_{i1}, \dots, w_{in})$ ,  $i = 1, 2$ , and let  $\Sigma_Z$  be a covariance matrix of the vector  $Z = (Z_{11}, \dots, Z_{1n}, Z_{21}, \dots, Z_{2n})$  as defined in (2). One can show that the joint pdf for vector  $\mathbf{W}^* = (W_{11}^*, \dots, W_{1n}^*, W_{21}^*, \dots, W_{2n}^*)$ , where  $W_{ij}^* = Z_{ij} + \alpha_{i0}^U \mathcal{E}_0^U - \alpha_{i0}^L \mathcal{E}_0^L$ , is

$$f_{n,2}^{\mathbf{W}^*}(\mathbf{w}_1, \mathbf{w}_2) = K_w \exp \left\{ \frac{c_1^2 c_{22} + 2c_1 c_2 c_{12} + c_2^2 c_{11}}{2c_\delta} \right\} \Phi_{\rho^*} \left( \frac{c_1 c_{22} + c_2 c_{12}}{(c_\delta c_{22})^{1/2}}, \frac{c_1 c_{12} + c_2 c_{11}}{(c_\delta c_{11})^{1/2}} \right),$$

where  $c_\delta = c_{11}c_{22} - c_{12}^2$ ,  $K_w = (2\pi)^{1-n} (c_\delta \det(\Sigma))^{-1/2} \exp\{-\mathbf{w}^T \Sigma^{-1} \mathbf{w}/2\}$ ,  $\Phi_{\rho^*}$  is the cdf of a bivariate standard normal random variable with correlation  $\rho^*$  and

$$c_1 = c_1(\mathbf{w}) = \alpha_{10}^U s_1(\mathbf{w}) + \alpha_1^U s_2(\mathbf{w}) - 1, \quad c_2 = c_2(\mathbf{w}) = -\alpha_{10}^L s_1(\mathbf{w}) - \alpha_1^L s_2(\mathbf{w}) - 1,$$

$$c_{11} = (\alpha_{10}^U)^2 s_{11} + 2\alpha_{10}^U \alpha_1^U s_{12} + (\alpha_1^U)^2 s_{22}, \quad c_{22} = (\alpha_{10}^L)^2 s_{11} + 2\alpha_{10}^L \alpha_1^L s_{12} + (\alpha_1^L)^2 s_{22},$$

$$c_{12} = \alpha_{10}^U \alpha_{10}^L s_{11} + (\alpha_{10}^U \alpha_1^L + \alpha_{10}^L \alpha_1^U) s_{12} + \alpha_1^U \alpha_1^L s_{22},$$

$$s_1 = s_1(\mathbf{w}) = \sum_{j=1}^n (\Sigma_Z^{-1} \mathbf{w})_j, \quad s_2 = s_2(\mathbf{w}) = \sum_{j=n+1}^{2n} (\Sigma_Z^{-1} \mathbf{w})_j,$$

$$s_{11} = \sum_{j_1, j_2=1}^n (\Sigma_Z^{-1})_{j_1, j_2}, \quad s_{22} = \sum_{j_1, j_2=n+1}^{2n} (\Sigma_Z^{-1})_{j_1, j_2}, \quad s_{12} = \sum_{j_1=1}^n \sum_{j_2=n+1}^{2n} (\Sigma_Z^{-1})_{j_1, j_2}.$$

The pdf of  $V_i = \alpha_i^U \mathcal{E}_0^U - \alpha_i^L \mathcal{E}_0^L$  is  $f_{V_i}(w) = \{\exp\{-w_+/\alpha_i^U - (-w)_+/\alpha_i^L\}/(\alpha_i^U + \alpha_i^L)\}$ . We use the convolution formula to get

$$f_{n,2}^{\mathbf{W}}(\mathbf{w}_1, \mathbf{w}_2) = \int_{\mathbb{R}^2} f_{n,2}^{\mathbf{W}^*}(\mathbf{w}_1 - v_1, \mathbf{w}_2 - v_2) f_{V_1}(v_1) f_{V_2}(v_2) dv_1 dv_2.$$

If we assume that  $\alpha_2^U = \alpha_2^L = 0$ , the formula simplifies to a one-dimensional integral:

$$\begin{aligned} f_{n,2}^{\mathbf{W}}(\mathbf{w}_1, \mathbf{w}_2) &= \int_{\mathbb{R}^1} f_{n,2}^{\mathbf{W}^*}(\mathbf{w}_1 - v_1, \mathbf{w}_2) f_{V_1}(v_1) dv_1 \\ &= \frac{1}{\alpha_1^U + \alpha_1^L} \int_{\mathbb{R}_+} \{f_{n,2}^{\mathbf{W}^*}(\mathbf{w}_1 - v_1, \mathbf{w}_2) \exp(-v_1/\alpha_1^U) + f_{n,2}^{\mathbf{W}^*}(\mathbf{w}_1 + v_1, \mathbf{w}_2) \exp(-v_1/\alpha_1^L)\} dv_1. \end{aligned}$$

This integral can be evaluated with very good accuracy via Gauss-Legendre quadrature using 25 – 30 quadrature points; see Stroud and Secrest (1966) for details.

## A.2 Proof of Proposition 1

For simplicity, we omit indices for the correlation coefficient  $\rho = \rho_{1,2}^{1;2}$ . We have:

$$F_{2,1:2}^W(z_1, z_2) = \int_{\mathbb{R}_+^3} \Phi_\rho(z_1 - \alpha_{10}^U v_0 - \alpha_1^U v_1, z_2 - \alpha_{20}^U v_0 - \alpha_2^U v_2) \exp(-v_0 - v_1 - v_2) dv_0 dv_1 dv_2.$$

We use the integration by parts formula with respect to  $v_1$  and  $v_2$  to find that

$$F_{2,1:2}^W(z_1, z_2) = \int_{\mathbb{R}_+^1} \{I_0(v_0) - I_1(v_0) - I_2(v_0) + I_{12}(v_0)\} dv_0, \quad (8)$$

where  $I_0(v_0) = \Phi_\rho(z_1 - \alpha_{10}^U v_0, z_2 - \alpha_{20}^U v_0) \exp(-v_0)$ ,  $I_1(v_0) = \Phi_\rho(z_1 - \alpha_{10}^U v_0 - \rho/\alpha_2^U, z_2 - \alpha_{20}^U v_0 - 1/\alpha_2^U) \exp\{(\delta_2 - 1)v_0 + 0.5/(\alpha_2^U)^2 - z_2/\alpha_2^U\}$ ,  $I_2(v_0) = \Phi_\rho(z_1 - \alpha_{10}^U v_0 - 1/\alpha_1^U, z_2 - \alpha_{20}^U v_0 - \rho/\alpha_1^U) \exp\{(\delta_1 - 1)v_0 + 0.5/(\alpha_1^U)^2 - z_1/\alpha_1^U\}$ ,  $I_{12}(v_0) = \Phi_\rho(z_1 - \alpha_{10}^U v_0 - 1/\alpha_1^U - \rho/\alpha_2^U, z_2 - \rho/\alpha_2^U - \alpha_{20}^U v_0 - 1/\alpha_2^U) \exp\{(\delta_{12} - 1)v_0 + 0.5(\rho_{12}^*)^2 - z_1/\alpha_1^U - z_2/\alpha_2^U\}$ , and  $(\rho_{12}^*)^2 = \{(\alpha_1^U)^2 + 2\rho\alpha_1^U\alpha_2^U + (\alpha_2^U)^2\}/(\alpha_1^U\alpha_2^U)^2$ .

For the marginal distribution,  $F_{1,i}^W(z) = \Phi(z) - [\alpha_{i0}^U \exp\{-z/\alpha_{i0}^U + 0.5/(\alpha_{i0}^U)^2\} \Phi(z - 1/\alpha_{i0}^U) - \alpha_i^U \exp\{-z/\alpha_i^U + 0.5/(\alpha_i^U)^2\} \Phi(z - 1/\alpha_i^U)]/(\alpha_{i0}^U - \alpha_i^U)$ ,  $i = 1, 2$ . Let  $\delta_i \neq 1$ ,  $\alpha_i^* = \max(\alpha_{i0}^U, \alpha_i^U)$  and let  $z_i = c_i + \alpha_i^* \log n$ , where  $c_i = 0.5/\alpha_i^* + \alpha_i^* \log(\alpha_i^*/|\alpha_{i0}^U - \alpha_i^U|) - \alpha_i^* \log x_i$ . This implies that  $F_{1,i}^W(z_i) = 1 - x_i/n + o(1/n)$ .

*Case 1:*  $\delta_1 > 1, \delta_2 > 1$ . We get

$$\begin{aligned} \int_{\mathbb{R}_+^1} I_0(v_0) dv_0 &= \int_{\mathbb{R}_+^1} \Phi_\rho\{c_1 - \alpha_{10}^U(v_0 - \log n), c_2 - \alpha_{20}^U(v_0 - \log n)\} \exp(-v_0) dv_0 \\ &= 1 - \frac{1}{n} \sum_{k=1}^2 \alpha_{k0}^U \int_{\mathbb{R}^1} \Phi \left\{ \frac{c_{3-k} - \rho c_k + (\alpha_{3-k,0}^U - \rho \alpha_{k0}^U) v}{(1 - \rho^2)^{1/2}} \right\} \phi(c_k + \alpha_{k0}^U v) \exp(v) dv + o\left(\frac{1}{n}\right). \end{aligned} \quad (9)$$



To compute the integrals in (9), we use the following equality:

$$\int_{\mathbb{R}^1} \exp(\theta v) \phi(v) \Phi(qv) dv = \exp(0.5\theta^2) \Phi \left\{ \frac{\theta q}{(q^2 + 1)^{1/2}} \right\}. \quad (10)$$

This equality can be obtained by differentiating the integral on the left-hand side with respect to the parameter  $q$ . We apply (10) to (9) and, after combining all terms, we get:

$$\int_{\mathbb{R}_+^1} I_0(v_0) dv_0 = 1 - \frac{y_1}{n} \Phi \left\{ \frac{\rho_{12}}{2} + \frac{\log(y_1/y_2)}{\rho_{12}} \right\} - \frac{y_2}{n} \Phi \left\{ \frac{\rho_{12}}{2} + \frac{\log(y_2/y_1)}{\rho_{12}} \right\} + o \left( \frac{1}{n} \right).$$

We use (10) to compute other terms in (8). Let  $c_1^* = c_1 - \frac{1}{\alpha_1^U} - \frac{\rho}{\alpha_2^U}$  and  $c_2^* = c_2 - \frac{\rho}{\alpha_1^U} - \frac{1}{\alpha_2^U}$ , and let  $c_{12} = \exp \left\{ -\frac{c_1}{\alpha_1^U} - \frac{c_2}{\alpha_2^U} + 0.5(\rho_{12}^*)^2 \right\} = \exp \left\{ -\frac{c_1^*}{\alpha_1^U} - \frac{c_2^*}{\alpha_2^U} \right\}$ . We get

$$\begin{aligned} \int_{\mathbb{R}_+^1} I_{12}(v_0) dv_0 &= \frac{c_{12}}{n} \int_{\mathbb{R}^1} \Phi(\alpha_{10}^U v + c_1^*, \alpha_{20}^U v + c_2^*) \exp\{-(\delta_{12} - 1)v\} dv + o \left( \frac{1}{n} \right) \\ &= \frac{c_{12}}{n} \sum_{k=1}^2 \alpha_{k0}^U \int_{\mathbb{R}^1} \Phi \left\{ \frac{c_{3-k}^* - \rho c_k^* + (\alpha_{3-k,0}^U - \rho \alpha_{k0}^U) v}{(1 - \rho^2)^{1/2}} \right\} \phi(\alpha_{k0}^U v + c_k^*) \exp\{-(\delta_{12} - 1)v\} dv + o \left( \frac{1}{n} \right) \\ &= \frac{1}{n} (\delta_{12} - 1)^{-1} y_1^{\delta_1} y_2^{1-\delta_1} \delta_1^* \exp\{0.5\delta_1(\delta_1 - 1)\rho_{12}^2\} \Phi \left\{ \rho_{12}(0.5 - \delta_1) + \frac{\log(y_2/y_1)}{\rho_{12}} \right\} \\ &\quad + \frac{1}{n} (\delta_{12} - 1)^{-1} y_2^{\delta_2} y_1^{1-\delta_2} \delta_2^* \exp\{0.5\delta_2(\delta_2 - 1)\rho_{12}^2\} \Phi \left\{ \rho_{12}(0.5 - \delta_2) + \frac{\log(y_1/y_2)}{\rho_{12}} \right\} + o \left( \frac{1}{n} \right). \end{aligned}$$

Similarly, we can calculate the two remaining terms in (8), and, after combining terms and taking limit as  $n \rightarrow \infty$ , we can show that (4) holds.

*Case 2:*  $\delta_1 > 1, \delta_2 < 1$ . Denote  $\alpha_2^* = \alpha_2^U - \alpha_{20}^U > 0$ . We get:

$$\begin{aligned} \int_{\mathbb{R}_+^1} I_0(v_0) dv_0 &= \int_{\mathbb{R}_+^1} \Phi_\rho \{c_1 - \alpha_{10}^U(v_0 - \log n), c_2 - \alpha_{20}^U(v_0 - \log n) + \alpha_2^* \log n\} \exp(-v_0) dv_0 \\ &= \frac{1}{n} \int_{-\log n}^\infty \Phi(c_1 - \alpha_{10}^U v) \exp(-v) dv + o \left( \frac{1}{n} \right) = 1 - \frac{1}{n} \exp \left\{ \frac{0.5}{(\alpha_{10}^U)^2} - \frac{c_1}{\alpha_{10}^U} \right\} + o \left( \frac{1}{n} \right) \\ &= 1 - \frac{x_1}{n} \left( 1 - \frac{1}{\delta_1} \right) + o \left( \frac{1}{n} \right), \quad (11) \end{aligned}$$

$$\begin{aligned} \int_{\mathbb{R}_+^1} I_{12}(v_0) dv_0 &= \frac{c_{12}}{n^{1+\alpha_2^*/\alpha_2^U}} \int_{-\infty}^{\log n} \Phi(\alpha_{10}^U v + c_1^*, \alpha_{20}^U v + c_2^* + \alpha_2^* \log n) \exp\{-(\delta_{12} - 1)v\} dv \\ &= \frac{c_{12}}{n^{1+\alpha_2^*/\alpha_2^U}} \int_{\mathbb{R}^1} \Phi(\alpha_{10}^U v + c_1^*) \exp\{-(\delta_{12} - 1)v\} dv + o \left( \frac{1}{n} \right) = o \left( \frac{1}{n} \right). \quad (12) \end{aligned}$$

Similarly, we find that

$$\int_{\mathbb{R}_+^2} I_1(v_0) dv_0 = \frac{x_2}{n} + o\left(\frac{1}{n}\right), \quad \int_{\mathbb{R}_+^2} I_1(v_0) dv_0 = \frac{x_1}{n\delta_1} + o\left(\frac{1}{n}\right).$$

Therefore,  $\ell(x_1, x_2) = x_1 + x_2$ . This implies that  $C_{2,1:2}^W$  as well as the limiting extreme-value copula  $C_{2,1:2}^W$  has no upper tail dependence. The remaining two cases when  $\delta_1 < 1, \delta_2 > 1$  and when  $\delta_1 < 1, \delta_2 < 1$  are considered analogously.  $\square$

### A.3 Proof of Proposition 2

For  $c_i \rightarrow \infty$ , we use the asymptotic property of the sum of two independent Pareto variables with regular varying tails (see Feller (1970)) to get:

$$\begin{aligned} \text{pr}\{W_{i1} = Z_{i1} + \alpha_{i0}^U \mathcal{E}_0 + \alpha_i^U \mathcal{E}_i < c_i\} &= \text{pr}\{\alpha_{i0}^U \mathcal{E}_0 + \alpha_i^U \mathcal{E}_i < c_i\} + o(c_i^{-k}) \\ &= 1 - \{(\alpha_{i0}^U)^k + (\alpha_i^U)^k\} c_i^{-k} + o(c_i^{-k}), \quad i = 1, 2. \end{aligned}$$

Let  $c_i := (n/x_i)^{1/k} \cdot \{(\alpha_{i0}^U)^k + (\alpha_i^U)^k\}^{1/k}$ . This implies  $\text{pr}\{W_{i1} < c_i\} = 1 - x_i/n + o(1/n)$ ,  $i = 1, 2$ . We have

$$\begin{aligned} \text{pr}\{W_{11} < c_1, W_{21} < c_2\} &= \text{pr}\{\alpha_{10}^U \mathcal{E}_0 + \alpha_1^U \mathcal{E}_1 < c_1, \alpha_{20}^U \mathcal{E}_0 + \alpha_2^U \mathcal{E}_2 < c_2\} + o\left(\frac{1}{n}\right) \\ &= k \int_1^{\mathcal{Q}} \left\{ 1 - \left( \frac{c_1 - \alpha_{10}^U v_0}{\alpha_1^U} \right)^{-k} \right\} \left\{ 1 - \left( \frac{c_2 - \alpha_{20}^U v_0}{\alpha_2^U} \right)^{-k} \right\} v_0^{-k-1} dv_0 + o\left(\frac{1}{n}\right), \end{aligned}$$

where  $\mathcal{Q} = \min\{(c_1 - \alpha_1^U)/\alpha_{10}^U, (c_2 - \alpha_2^U)/\alpha_{20}^U\}$ . Let  $x_1\theta_1^* > x_2\theta_2^*$  so that  $\mathcal{Q} = (c_1 - \alpha_1^U)/\alpha_{10}^U$  as  $n \rightarrow \infty$ . We find that  $\text{pr}\{W_{11} < c_1, W_{21} < c_2\} = I_1 + I_2 + I_{12} + o(1/n)$ , where

$$\begin{aligned} S_1 &= k \int_1^{\mathcal{Q}} \left\{ 1 - \left( \frac{c_1 - \alpha_{10}^U v_0}{\alpha_1^U} \right)^{-k} \right\} v_0^{-k-1} dv_0 = \text{pr}\{\alpha_{10}^U \mathcal{E}_0 + \alpha_1^U \mathcal{E}_1 < c_1\} = 1 - x_1/n + o\left(\frac{1}{n}\right), \\ S_2 &= -k \int_1^{\mathcal{Q}} \left( \frac{c_2 - \alpha_{20}^U v_0}{\alpha_2^U} \right)^{-k} v_0^{-k-1} dv_0 = (\delta^*)^{-k} k \int_1^{\mathcal{Q}} \left\{ 1 - \left( \frac{c_2 - \alpha_{20}^U v_0}{\delta^* \alpha_2^U} \right)^{-k} \right\} v_0^{-k-1} dv_0 \end{aligned}$$

$$\begin{aligned}
-(\delta^*)^{-k} \int_1^{\mathcal{Q}} v_0^{-k-1} dv_0 &= (\delta^*)^{-k} \Pr\{\alpha_{20}^U \mathcal{E}_0 + \delta^* \alpha_2^U \mathcal{E}_2 < c_2\} - (\delta^*)^{-k} \left\{ 1 - \left( \frac{c_1 - \alpha_1^U}{\alpha_{10}^U} \right)^{-k} \right\} \\
&= (\delta^*)^{-k} \left[ \left( \frac{c_1 - \alpha_1^U}{\alpha_{10}^U} \right)^{-k} - c_2^{-k} \{(\alpha_{20}^U)^k + (\delta^* \alpha_2^U)^k\} \right] + o\left(\frac{1}{n}\right),
\end{aligned}$$

where  $\delta^* = (c_2 - \alpha_{20}^U \mathcal{Q})/\alpha_2^U$ . It follows that

$$S_2 = \frac{x_1}{n} \cdot \frac{(\alpha_{10}^U/\delta^*)^k}{(\alpha_{10}^U)^k + (\alpha_1^U)^k} - \frac{x_2}{n} \cdot \frac{(\alpha_{20}^U/\delta^*)^k + (\alpha_2^U)^k}{(\alpha_{20}^U)^k + (\alpha_2^U)^k} + o\left(\frac{1}{n}\right) = -\frac{x_2}{n} \cdot \frac{(\alpha_2^U)^k}{(\alpha_{20}^U)^k + (\alpha_2^U)^k} + o\left(\frac{1}{n}\right).$$

Lastly,  $S_{12} = k \int_1^{\mathcal{Q}} \left( \frac{c_1 - \alpha_{10}^U v_0}{\alpha_1^U} \right)^{-k} \left( \frac{c_2 - \alpha_{20}^U v_0}{\alpha_2^U} \right)^{-k} v_0^{-k-1} dv_0 = o\left(\frac{1}{n}\right)$ . Therefore,  $\ell(x_1, x_2) = x_1 + (\alpha_2^U)^k x_2 / \{(\alpha_{20}^U)^k + (\alpha_2^U)^k\} = x_1 + x_2 - \theta_2^* x_2$ . Similarly, if  $\theta_1^* x_1 < \theta_2^* x_2$ , we get  $\ell(x_1, x_2) = x_1 + x_2 - \theta_1^* x_1$  so that  $\ell(x_1, x_2) = x_1 + x_2 - \min(\theta_1^* x_1, \theta_2^* x_2)$ .  $\square$

## A.4 Proof of Proposition 3

We have  $\lambda_U = \lim_{n \rightarrow \infty} n \Pr(W_{11} > c_1^n, W_{21} > c_2^n)$ , where  $F_{1,i}^W(c_i^n) = 1 - 1/n + o(1/n)$ . First, consider the case when  $\alpha_{i0}^U > \alpha_i^U$ ,  $i = 1, 2$ . By Embrechts et al. (1979),  $\Pr(W_{i1} > c_i^n) = \Pr(\alpha_{i0}^U \mathcal{E}_{i0}^U > c_i^n) + o(1/n) = \bar{F}_{\mathcal{E}}(c_i^n/\alpha_{i0}^U) + o(1/n)$ , where  $\bar{F}_{\mathcal{E}}(s) = 1 - F_{\mathcal{E}}(s)$ . This implies that  $c_i^n = \alpha_{i0}^U \bar{F}_{\mathcal{E}}^{-1}(1/n)$  can be selected and

$$\Pr(W_{11} > c_1^n, W_{21} > c_2^n) \leq \Pr(W_{11} > c_1^n) = 1/n + o(1/n),$$

$$\Pr(W_{11} > c_1^n, W_{21} > c_2^n) \geq \Pr\{\mathcal{E}_0^U + \min(Z_{11}/\alpha_{10}^U, Z_{21}/\alpha_{20}^U) > \bar{F}_{\mathcal{E}}^{-1}(1/n)\} = 1/n + o(1/n),$$

so that  $\lambda_U = 1$ .

Now let  $\alpha_{10}^U > \alpha_1^U$  and  $\alpha_{20}^U < \alpha_2^U$  (other cases are considered analogously). We get  $c_1^n = \alpha_{10}^U \bar{F}_{\mathcal{E}}^{-1}(1/n)$  and  $c_2^n = \alpha_2^U \bar{F}_{\mathcal{E}}^{-1}(1/n)$ . For any  $\epsilon > 0$ ,  $\Pr\{\mathcal{E}_0^U > (1+\epsilon)\bar{F}_{\mathcal{E}}^{-1}(1/n)\} = o(1/n)$  and therefore

$$\begin{aligned}
\Pr(W_{11} > c_1^n, W_{21} > c_2^n) &= \Pr\{W_{11} > c_1^n, W_{21} > c_2^n, Z_{21}/\alpha_{20}^U + \mathcal{E}_0^U < (1+\epsilon)\bar{F}_{\mathcal{E}}^{-1}(1/n)\} + o(1/n) \\
&\leq \Pr\{W_{11} > c_1^n, \alpha_2^U \mathcal{E}_2^U > \alpha_2^U \bar{F}_{\mathcal{E}}^{-1}(1/n) - \alpha_{20}^U(1+\epsilon)\bar{F}_{\mathcal{E}}^{-1}(1/n)\} + o(1/n)
\end{aligned}$$

$$= \text{pr}(W_{11} > c_1^n) \text{pr}[\alpha_2^U \mathcal{E}_2^U > \bar{F}_{\mathcal{E}}^{-1}(1/n) \{\alpha_2^U - \alpha_{20}^U(1 + \epsilon)\}] + o(1/n) = o(1/n)$$

for  $\epsilon > 0$  small enough such that  $\alpha_2^U - \alpha_{20}^U(1 + \epsilon) > 0$ . This implies  $\lambda_U = 0$ . □